

# Modifikasi Arsitektur VGG16 untuk Klasifikasi Citra Digital Rempah-Rempah Indonesia

## Classification of Indonesian Spices Digital Image using Modified VGG 16 Architecture

Evan Tanuwijaya<sup>1</sup>, Angelica Roseanne<sup>2</sup>  
Universitas Ciputra, Surabaya

### Article Info

#### Genesis Artikel:

Diterima, 7 Oktober 2021  
Direvisi, 21 Oktober 2021  
Disetujui, 22 Oktober 2021

#### Kata Kunci:

Convolution Neural Network  
Rempah-rempah  
VGG 16  
Alexnet  
Deep learning

### ABSTRAK

Rempah-rempah merupakan salah satu kekayaan alam yang dimiliki oleh Indonesia. Rempah-rempah sendiri memiliki banyak manfaat untuk Kesehatan ataupun hal-hal lain. Dari banyaknya rempah yang berada di Indonesia, ternyata masyarakat Indonesia sendiri masih memiliki pengetahuan yang rendah akan rempah-rempah tersebut. Hal ini menyebabkan banyak orang bahkan petani mengalami kesusahan dalam mengenali jenis rempah terutama remaja. Membedakan rempah satu dengan yang lain merupakan tantangan yang banyak dihadapi oleh masyarakat. Oleh sebab itu, penelitian ini membuat sebuah model klasifikasi dengan menggunakan *convolution neural network* dengan arsitektur VGG 16 yang dimodifikasi. Arsitektur modifikasi VGG 16 memiliki 10-layer yang terdiri dari 7-layer *convolution* dan 3-layer *fully connected*. Untuk fase latih model modifikasi VGG 16 ini menggunakan *dataset* rempah yang disediakan oleh Kaggle. Validasi model yang digunakan adalah akurasi, *loss*, *precision*, dan *recall* untuk membandingkan model mana yang memiliki nilai yang terbaik. Untuk model modifikasi VGG 16 yang dibuat untuk melakukan klasifikasi, mendapatkan hasil evaluasi rata-rata akurasi sebesar 81%, nilai *recall* sebesar 76%, dan nilai *precision* sebesar 81% untuk fase training dan untuk fase validasi, akurasi sebesar 85%, nilai *recall* sebesar 80%, dan nilai *precision* sebesar 84%. Jadi dengan model modifikasi VGG 16 dapat disimpulkan bahwa model mampu memprediksi rempah-rempah lebih baik dari model Alexnet.

### ABSTRACT

Spices are one of Indonesia's natural resources. Spices themselves have many benefits for health or other things. Of the many spices that are in Indonesia, it turns out that the Indonesian people themselves still have low knowledge of these spices. This causes many people and even farmers to have difficulty in recognizing the types of spices, especially teenagers. Distinguishing one spice from another is a challenge that many people face. Therefore, this study makes a classification model using a convolution neural network with a modified VGG 16 architecture. The modified VGG 16 architecture has a 10-layer consisting of 7-layer convolution and 3-layer fully connected. For the training phase, this modified VGG 16 model uses spice datasets provided by Kaggle. The model validation used is accuracy, loss, precision, and recall to compare which model has the best value. For the modified VGG 16 model which was made to perform classification, the evaluation results obtained an average accuracy of 81%, a recall value of 76%, and a precision value of 81% for the training phase and the validation phase, an accuracy of 85%, a recall value of 80%, and the precision value is 84%. So with the modified VGG 16 model, it can be concluded that the model can predict spices better than the Alexnet model.

This is an open access article under the [CC BY-SA](#) license.



### Penulis Korespondensi:

Evan Tanuwijaya,  
Program Studi Informatika,  
Universitas Ciputra Surabaya,  
Email: [evan.tanuwijaya@ciputra.ac.id](mailto:evan.tanuwijaya@ciputra.ac.id)

## 1. PENDAHULUAN

Indonesia merupakan negara yang banyak menghasilkan kekayaan alam, salah satunya adalah rempah - rempah. Rempah memiliki definisi nama yang bermacam - macam sesuai dengan tanaman asal maupun bentuknya. Rempah dapat berupa bunga, daun, biji, batang, dan akar yang dapat digunakan untuk menghasilkan aroma, memberikan warna, mencegah kerusakan, mengawetkan makanan, dan memiliki banyak sekali manfaat untuk menyehatkan tubuh [1], [2]. Seperti yang kita tahu, rempah di Indonesia sudah seakan menjadi keunikan tersendiri yang jarang ditemukan di negara lain. Salah satu manfaat utama dari rempah adalah sebagai bumbu dapur. Dari banyaknya rempah yang berada di Indonesia, ternyata masyarakat Indonesia sendiri masih memiliki pengetahuan yang rendah akan rempah-rempah tersebut. Seiring dengan perkembangan jaman dan budaya, pengetahuan akan jenis rempah semakin berkurang. Hal ini menyebabkan banyak orang bahkan petani mengalami kesusahan dalam mengenali jenis rempah terutama remaja [3]. Membedakan antara satu rempah dengan rempah yang lain merupakan suatu tantangan tersendiri bagi kalangan remaja akhir. Beberapa jenis rempah juga sekilas terlihat sama jika kita tidak mengetahui karakteristiknya. Dilansir dari Liputan6 [4], beberapa jenis rempah yang sulit dibedakan diantaranya yaitu jahe, lengkuas, kunyit dan kencur. Rempah-rempah tersebut memiliki karakteristik, bentuk, dan warna kulit yang hampir mirip dan susah untuk membedakan antara rempah satu dengan rempah yang lainnya. Sayangnya, pengenalan akan rempah kepada generasi muda masih sangat minim diajarkan oleh institusi pendidikan. Berdasarkan penelitian yang telah dilakukan, didapatkan hasil bahwa sebanyak 47% siswa di SMKN 9 Bandung masih belum mengenali bumbu dan rempah pada saat pengolahan makanan Indonesia [5].

Untuk membantu masyarakat dalam mengenali karakteristik dari rempah-rempah yang ada, penelitian untuk mendeteksi jenis-jenis rempah telah banyak dikembangkan. Dari penelitian yang dilakukan oleh Khrisne dan Suyadnya [6], mengusulkan sebuah metode menggunakan *Convolution Neural Network* untuk mengklasifikasikan rempah dari citra. Penelitian ini menggunakan metode CNN dengan arsitektur VGG yang di modifikasi menjadi small VGGNet-like Network. Model VGG yang digunakan ini dibuat dari model VGG yang dipotong beberapa bagian layer nya lalu di training untuk melakukan klasifikasi. Penelitian ini memiliki hasil akurasi sebesar 70% yang diuji menggunakan dataset dari Google Image Search. Penelitian selanjutnya diambil dari penelitian yang dilakukan oleh Isna Wulandari, Hasbi Yasin, dan Tatik [5]. Pada penelitian ini mengusulkan sebuah metode klasifikasi menggunakan CNN yang dibuat dari dua layer *convolution* di mana layer pertama dari *convolution* layer ini menggunakan jumlah filter sebanyak 10 dan ukuran dari filter yang digunakan adalah 3 x 3. Layer kedua dari CNN ini menggunakan jumlah layer sebanyak 20 layer. Hasil dari penelitian didapatkan akurasi sebesar 98% dengan nilai loss yaitu 0.076 yang didapatkan dari hasil *training* menggunakan data yang di *crawling* dari Google engine search. Penelitian selanjutnya diambil dari penelitian Kharuddin, Kusri, dan Emha [7]. Penelitian ini menggunakan metode *machine learning* yaitu K-nearest neighbour untuk melakukan klasifikasi dari fitur warna yang ada pada rempah. Dari citra rempah, diambil fitur warna RGB dan *texture color*, kemudian diproses menggunakan K-nearest neighbour untuk mengklasifikasikan jenis rempah. Hasil dari penelitian ini didapatkan akurasi sebesar 76% dari nilai K = 1.

Dari beberapa penelitian di atas, penelitian yang dilakukan [7] memiliki hasil yang cukup optimal dengan K-nearest neighbour menggunakan fitur RGB dan *texture-color* yaitu 76%. jika dibandingkan dengan penelitian yang dilakukan oleh [6] yang memiliki akurasi 70%, penelitian menggunakan K-nearest neighbour memiliki akurasi yang lebih tinggi dibandingkan dengan smaller VGGnet-like. Akan tetapi dengan menggunakan metode *convolution neural network*, dapat mempermudah proses ekstraksi fitur tanpa perlu mengetahui fitur-fitur apa saja yang digunakan karena sifat dari *convolution neural network* sendiri adalah *blackbox* [8]. Pada penelitian [5], hasil dari penelitian lebih baik dibandingkan dengan kedua penelitian yang lain. Arsitektur yang digunakan merupakan hasil rancangan sendiri dengan menggunakan 2 layer. Dan menghasilkan akurasi sebesar 98%. Berdasarkan dari hasil penelitian sebelumnya, pada penelitian ini akan dilakukan klasifikasi citra digital rempah-rempah Indonesia menggunakan modifikasi arsitektur VGG16. Arsitektur VGG 16 [9] merupakan hasil pengembangan dari Alexnet [10] di mana arsitektur ini berfokus pada memperbanyak proses fitur ekstraksi pada layer *convolution* sehingga mampu mendapatkan representasi citra yang banyak untuk dapat di klasifikasikan. Penelitian ini akan menggunakan arsitektur VGG16 yang dimodifikasi dengan mengurangi jumlah layer dan jumlah *kernel* yang ada pada masing-masing layer untuk mencapai hasil yang optimal. Penelitian ini menggunakan dataset yang disediakan oleh Kaggle [11] untuk melakukan *training* dan *testing* untuk model modifikasi VGG16 yang dibuat. Untuk penelitian ini, peneliti melakukan perbedaan pada model yang digunakan dengan model dasar yaitu VGG16 yang memiliki arsitektur lebih sedikit dibandingkan dengan VGGnet-Like dan membandingkan nya dengan beberapa arsitektur state-of-the-art. Oleh sebab itu, pada penelitian ini akan dijelaskan pada 4 bagian yaitu bagian 1 untuk pendahuluan, bagian 2 untuk menjelaskan metode penelitian, bagian 3 untuk menjelaskan hasil dan Analisa dari penelitian, dan terakhir bagian 4 menjelaskan kesimpulan dari penelitian ini.

## 2. METODE PENELITIAN

Pada bagian ini, akan dijelaskan terkait dengan metode apa saja yang akan digunakan, penjelasan terkait dataset yang akan digunakan, proses yang akan dilakukan, dan evaluasi yang akan digunakan.

### 2.1. Datasets

Dataset yang digunakan diambil dari Kaggle [11]. Dataset tersebut berupa citra digital rempah yang terdiri atas folder *train* dan *test*. Terdapat 100 data *train* dan 25 data *test* yang terdiri termasuk ke dalam lima jenis rempah yaitu jahe, kencur, kunyit,

lengkuas, dan temulawak. Masing-masing citra memiliki ukuran sebesar  $2354 \times 1773$ . Citra rempah dapat dilihat pada Gambar 1. Citra yang didapatkan ini diambil dengan latar belakang putih untuk mempermudah dalam memisahkan antara objek dan latar belakang.



Gambar 1. Contoh dataset rempah

## 2.2. Preprocessing Data

Pada tahapan *preprocessing*, dilakukan augmentasi dataset dengan menggunakan *library* Image Data Generator dari Tensorflow dengan parameter sebagai berikut:

1. *Rescale* =  $1/255$
2. *Rotation Range* = 30
3. *Zoom Range* = 0.2
4. *Horizontal Flip* = true
5. *Vertical Flip* = true

Nilai dari masing-masing pixel pada gambar di skala menjadi  $1/255$ , kemudian gambar akan secara acak di putar dengan maksimal 30 derajat, Kemudian akan dilakukan *zoom*, *flip* baik itu secara horizontal ataupun vertikal. Proses augmentasi ini akan diterapkan pada data training. Untuk ukuran gambar diubah menjadi  $224 \times 224$  untuk setiap data *training* dan *testing* karena model dari Alexnet sendiri menggunakan ukuran  $224 \times 224$ .

## 2.3. Model Convolution Neural Network

Penelitian ini akan memodifikasi model VGG 16 yang dimodifikasi beberapa layernya yaitu mengurangi jumlah layer dan parameternya. VGG 16 [12] memiliki total layer 16 yaitu 13 layer *convolution* dan 3 layer *fully connected*. VGG16 menggunakan konsep blok untuk membentuk *Convolution* layer yang masing-masing *convolution* layer nya memiliki ukuran  $3 \times 3$  dan *stride* 1. Pada akhir blok, digunakan *max pooling* layer dengan ukuran  $2 \times 2$  dan *stride* 2. Dari arsitektur VGG 16 tersebut, peneliti memodifikasi dengan menggunakan konsep VGG dan menghasilkan model *convolution neural network* dengan arsitektur VGG 16 yang di modifikasi seperti tabel 1

Tabel 1. Arsitektur Modifikasi VGG 16

Name	Filter size	Stride	Jumlah
Input	$224 \times 224 \times 3$		
Convolution 1	$3 \times 3$	1	16
Convolution 2	$3 \times 3$	1	16
Maxpool 1	$2 \times 2$	2	16
Convolution 3	$3 \times 3$	1	32
Convolution 4	$3 \times 3$	1	32
Maxpool 2	$2 \times 2$	2	32
Convolution 5	$3 \times 3$	1	64
Convolution 6	$3 \times 3$	1	64
Convolution 7	$3 \times 3$	1	64
Maxpool 3	$2 \times 2$	2	64
FC 1			1000
FC 2			700
Output			5

Tabel 2. Arsitektur Alexnet

Name	Filter size	Stride	Jumlah
<b>Input</b>	224 × 224 × 3		
<b>Convolution 1</b>	11 × 11	4	96
<b>Maxpool 1</b>	3 × 3	2	96
<b>Convolution 2</b>	5 × 5	1	256
<b>Maxpool 2</b>	3 × 3	2	256
<b>Convolution 3</b>	3 × 3	1	384
<b>Convolution 4</b>	3 × 3	1	384
<b>Convolution 5</b>	3 × 3	1	256
<b>Maxpool 3</b>	3 × 3	2	256
<b>FC 1</b>			4096
<b>FC 2</b>			4096
<b>Output</b>			5

Pada modifikasi VGG 16 [13]–[15], arsitektur yang digunakan berjumlah 10 layer. Seperti pada Tabel 1, ukuran filter yang digunakan adalah 3 × 3 dan filter pada *max pooling* adalah 2 × 2. Untuk *stride* yang digunakan adalah 1 pada layer *convolution* dan 2 pada layer *max pooling*. Untuk jumlah *kernel* pada masing-masing layer dikurangi yaitu pada layer 1 dan 2 menggunakan jumlah 16 sehingga *maxpooling*-nya berjumlah 16. Pada layer 3 dan 4 menggunakan jumlah 32, dan pada layer 5, 6, 7 menggunakan jumlah 64. Hasil dari *convolution* layer kemudian akan di klasifikasikan menggunakan *fully connected* layer sebanyak 2 buah layer yang masing-masing layer nya memiliki jumlah neuron 1000 dan 700.

Untuk menguji model modifikasi VGG 16, akan digunakan model arsitektur Alexnet [16], [17] dengan arsitektur seperti pada tabel 2. Alexnet memiliki 8 layer yaitu 5 layer *convolution* dan 3 layer *fully connected* layer. Untuk layer pertama dari Alexnet memiliki jumlah filter sebanyak 96 dan ukuran 11 × 11 dan *stride* 4 kemudian diikuti dengan *maxpooling* dengan ukuran 3 × 3 dan *stride* 2. Untuk layer *convolution* kedua memiliki ukuran 5 × 5 dan *stride* 1 sejumlah 256 diikuti oleh *maxpooling* dengan ukuran 3 × 3 dan *stride* 2. Untuk layer 3 dan 4 menggunakan ukuran 3 × 3 dengan *stride* 1 dan jumlah layer sebanyak 384 pada masing-masing layer-nya. Untuk layer 3 dan 4 tidak menggunakan *maxpooling*. Pada *convolution* layer ke 5 digunakan ukuran 3 × 3 dan *stride* 1 dengan jumlah *kernel* yaitu 256 diikuti dengan *maxpooling* dengan ukuran 3 × 3 dengan *stride* 2. Setelah di proses dengan *convolution* layer, *fully connected* layer akan mengklasifikasikan hasil dari *convolution* layer. Untuk layer pertama dan kedua digunakan 4096 neuron.

Pada kedua model yang digunakan, memiliki *input* layer yaitu 224 × 224. Ukuran ini merupakan ukuran asli dari masing-masing model. Untuk angka 3 pada masing-masing ukuran melambangkan *channel* dari warna citra yaitu *red*, *green* dan *blue* (RGB). Untuk *output* layer menggunakan nilai 5 karena pada penelitian ini, dataset yang digunakan memiliki 5 buah kelas yaitu: jahe, kencur, kunyit, lengkuas dan temulawak

### 3. HASIL DAN ANALISIS

Pada bagian ini akan di bahas mengenai hasil dari pelatihan model modifikasi VGG16 dan Alexnet. Untuk fase pelatihan model, peneliti menggunakan Bahasa pemrograman Python untuk membuat model dan dibantu dengan *library* Tensorflow yang merupakan salah satu *library* Python yang sangat terkenal untuk membuat model *Deep Learning*. Pembuatan model dilakukan dengan IDE Jupyter Notebook pada Google Collaboratory. Pada Google Collaboratory disediakan GPU yang dapat digunakan oleh pengguna secara gratis dengan beberapa batasan. Akan tetapi untuk penelitian ini, penggunaan Google Collaboratory untuk membuat penelitian sudah cukup. Pembuatan model menggunakan Tensorflow ini memiliki beberapa parameter yang perlu diterapkan. Kedua mode menggunakan parameter pelatihan yang sama untuk mengetahui model mana yang lebih optimal dari kedua model tersebut. Untuk parameter yang digunakan adalah parameter *optimizer* yang digunakan adalah adam [18] dengan *learning rate* 0,001. Untuk fungsi *loss* yang digunakan adalah *categorical crossentropy*. Fase *training* menggunakan *epoch* sebesar 100 dan langsung menggunakan *validation* dataset untuk melihat hasil dari model tersebut. Untuk evaluasi model digunakan akurasi, *precision*, dan *recall* untuk membandingkan kedua model tersebut.

Untuk model modifikasi VGG 16 memiliki total parameter sebanyak 50.990.223 dan untuk Alexnet memiliki total parameter sebanyak 228.436.037. Dari total jumlah parameter, modifikasi VGG 16 memiliki jumlah parameter yang jauh lebih sedikit dibandingkan dengan Alexnet. Jumlah parameter yang sedikit akan mengurangi total durasi *training* dari model dan mempercepat proses prediksi. Untuk mengevaluasi kedua model yang latih, peneliti menggunakan akurasi, *loss*, *precision*, dan *recall* untuk membandingkan model mana yang memiliki nilai yang terbaik. Akurasi dihitung dengan mencari nilai *true positif*, *true negatif*, *false positif*, dan *false negative* antara kelas dan hasil prediksi dari kedua model seperti pada rumus (1)

$$Akurasi = \frac{(tp+tn)}{(tp+tn+fp+fn)} \quad (1)$$

Untuk menghitung nilai *precision*, digunakan nilai *true positif* dibagi dengan nilai *true positif* ditambah dengan nilai *false positif* dimana *false positif* adalah data kelas yang salah di prediksi sebagai benar. Penghitungan *precision* dapat dilihat pada rumus (2).

$$precision = \frac{tp}{(tp+fp)} \quad (2)$$

Nilai *recall* dihitung dari nilai *true positif* dibagi dengan nilai *true positif* di jumlah dengan nilai *true negative* dimana *true negative* adalah nilai kelas yang benar di prediksi sebagai salah. Penghitungan *recall* dapat dilihat pada rumus (3).

$$recall = \frac{tp}{(tp+fn)} \quad (3)$$

Evaluasi model menggunakan akurasi, *precision*, dan *recall* dapat dilihat pada Tabel 3. Pada tabel tersebut, perbandingan 2 model di evaluasi dari hasil terakhir fase *training* dan rata-rata dari 100 *epoch*. Durasi dari *training* antara model modifikasi VGG 16 dan Alexnet tidak jauh berbeda yaitu 11 menit dan memiliki selisih 24 detik. Dari durasi *training* ini Modifikasi VGG 16 lebih singkat dikarenakan total parameter yang jumlahnya jauh lebih sedikit dibandingkan dengan Alexnet. Untuk nilai *loss* didapatkan model modifikasi VGG 16 lebih rendah dibandingkan dengan Alexnet baik itu pada epoch terakhir maupun rata-rata dari keseluruhan fase *training*. Modifikasi VGG 16 memiliki nilai *loss* sebesar 0,032 pada epoch terakhir dan rata-rata keseluruhan *training* sebesar 0,446. Sementara Alexnet memiliki nilai *loss* sebesar 0,645 pada epoch terakhir dan rata-rata *loss* sebesar 3,586. Modifikasi VGG 16 memiliki nilai akurasi sebesar 0,99, nilai *recall* sebesar 0,99, dan nilai *precision* sebesar 0,99 pada epoch terakhir dan rata-rata fase *training* akurasi sebesar 0,815, nilai *recall* sebesar 0,767 dan nilai *precision* sebesar 0,818. Jika hasil evaluasi model modifikasi VGG 16 dibandingkan dengan hasil evaluasi dari Alexnet, maka dapat dilihat bahwa pada epoch terakhir model modifikasi VGG 16 memiliki nilai yang lebih tinggi dibandingkan dengan Alexnet. Pada perbandingan nilai rata-rata *training*, dapat dilihat bahwa nilai *recall* dari Alexnet memiliki nilai yang lebih tinggi dari pada modifikasi VGG 16 yaitu 0,791 untuk Alexnet dan 0,767 untuk modifikasi VGG 16.

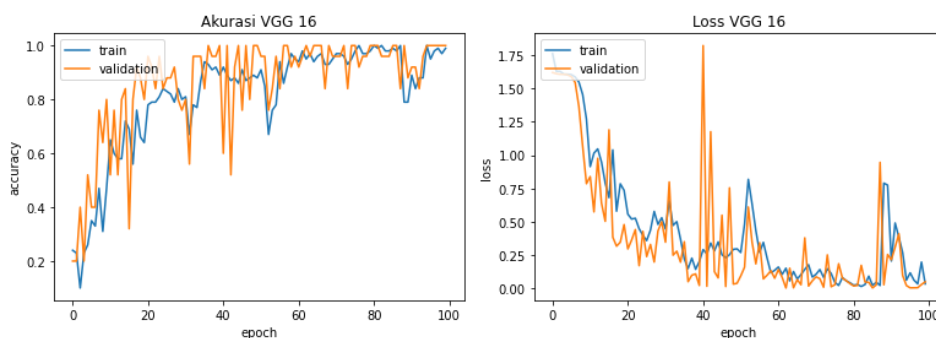
Validasi dari model dilakukan pada setiap *epoch* nya. Dapat dilihat pada tabel 3 bahwa untuk *loss* dari model modifikasi VGG16 memiliki nilai *loss* yang jauh lebih rendah yaitu 0,048 dibandingkan dengan nilai *loss* dari Alexnet yaitu 0,681 dan rata-rata nilai *loss* dari modifikasi VGG 16 adalah 0,376 jauh lebih rendah dibandingkan dengan Alexnet yang memiliki nilai 1155,570. Nilai akurasi, *precision*, dan *recall* dari hasil evaluasi model didapatkan modifikasi VGG 16 pada *epoch* terakhir adalah 1 dan untuk Alexnet memiliki nilai 0,8 untuk masing-masing akurasi, *recall*, dan *precision* nya. Rata-rata dari evaluasi model modifikasi VGG 16 memiliki nilai akurasi 0,857, nilai *recall* adalah 0,806, dan *precision* adalah 0,842 sementara untuk model alexnet jauh lebih rendah untuk rata-rata keseluruhan validasi pada masing-masing *epoch* yaitu nilai akurasi 0,331, nilai *recall* adalah 0,329, dan nilai *precision* adalah 0,333. Setiap *epoch* dapat dilihat nilai akurasi dan *loss* pada fase *training* dan validasi pada grafik yang dapat dilihat pada gambar 2 untuk modifikasi VGG 16 serta Gambar 3 untuk Alexnet.

Dari hasil evaluasi pada Tabel 3 didapatkan beberapa nilai terbaik banyak didapatkan oleh modifikasi VGG16.

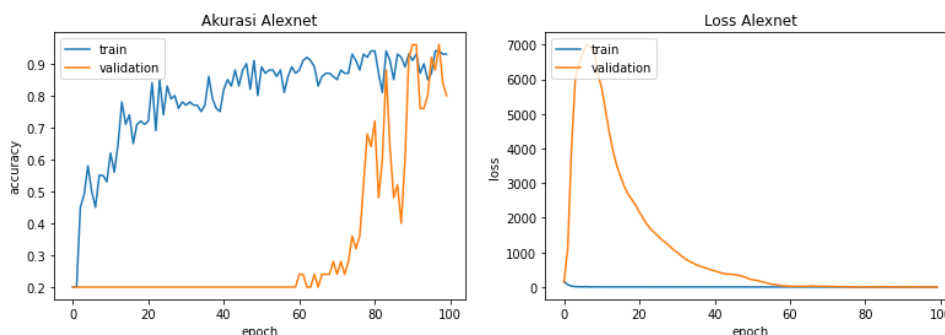
Tabel 3. Evaluasi Model Modifikasi VGG 16 dan Alexnet

	Modifikasi VGG 16		Alexnet	
	<i>Last</i>	<i>Average</i>	<i>Last</i>	<i>Average</i>
<b>Duration</b>	11.29		11.53	
<b>Loss</b>	0,032	0,446	0,645	3,586
<b>Akurasi</b>	0,990	0,815	0,930	0,799
<b>Recall</b>	0,990	0,767	0,930	0,791
<b>Precision</b>	0,990	0,818	0,930	0,808
<b>Val Loss</b>	0,048	0,376	0,681	1155,570
<b>Val Akurasi</b>	1,000	0,857	0,800	0,331
<b>Val recall</b>	1,000	0,806	0,800	0,329
<b>Val Precision</b>	1,000	0,842	0,800	0,333

Hal ini disebabkan oleh penggunaan ukuran filter dari VGG 16 yang lebih kecil pada layer pertama dan kedua menyebabkan fitur yang didapatkan memiliki nilai informasi yang jauh lebih berguna dibandingkan dengan Alexnet yang menggunakan ukuran filter 11 x 11 pada layer pertama. Dapat dilihat pada Gambar 2 dan Gambar 3 bahwa grafik akurasi VGG 16 sudah mencapai



Gambar 2. Grafik Akurasi dan loss dari Model Modifikasi VGG 16



Gambar 3. Grafik Akurasi dan loss dari Model Alexnet

lebih besar dari 0,6 pada fase *training* dan validasi pada *epoch* ke 20 daripada Alexnet yang nilai validasinya baru naik pada *epoch* sekitar 60 dikarenakan Alexnet masih *overfitting* dan belum dapat menggeneralisasi data dengan baik. Jika dilihat dari jumlah parameter, modifikasi VGG 16 memiliki nilai parameter yang jauh lebih rendah dibandingkan dengan Alexnet sementara pada hasil evaluasi didapatkan modifikasi VGG 16 yang nilai parameternya lebih rendah memiliki nilai evaluasi yang lebih baik dibandingkan dengan Alexnet yang memiliki nilai parameter yang lebih banyak. Jika model modifikasi VGG ini dibandingkan dengan penelitian acuan [6], Untuk fase *training*, akurasi dari modifikasi VGG memiliki rata-rata akurasi sebesar 0,81 jika dibandingkan penelitian acuan yang akurasi *training*nya sebesar 0,78 pada *epoch* ke 100, hasil *training* modifikasi pada penelitian ini memiliki nilai yang lebih tinggi dan juga untuk validasi akurasi memiliki 0,52 pada penelitian acuan sementara pada penelitian ini, modifikasi VGG memiliki rata-rata akurasi sebesar 0,85. Hal ini terpengaruh oleh penggunaan ukuran filter yang lebih kecil untuk melakukan ekstraksi fitur pada citra rempah-rempah sehingga didapatkan fitur-fitur yang berguna untuk *fully connected* layer dapat mengklasifikasi rempah dengan baik.

#### 4. KESIMPULAN

Indonesia memiliki banyak sekali jenis rempah-rempah. Rempah-rempah yang ada memiliki banyak sekali kegunaan baik itu untuk aroma terapan dan Kesehatan. Akan tetapi, seiring dengan berkembangnya jaman, masyarakat semakin kesulitan dalam membedakan antara rempah yang satu dengan rempah yang lainnya. Dari banyaknya rempah yang berada di Indonesia, ternyata masyarakat Indonesia sendiri masih memiliki pengetahuan yang rendah akan rempah-rempah tersebut. Hal ini menyebabkan banyak orang bahkan petani mengalami kesusahan dalam mengenali jenis rempah terutama remaja. Untuk membantu masyarakat tersebut, model modifikasi VGG 16 dibuat untuk membantu dalam mengklasifikasi rempah-rempah. Penelitian ini memodifikasi model VGG 16 yang memiliki jumlah layer dan parameter yang lebih kecil dari pada penelitian sebelumnya dan menguji model tersebut memiliki akurasi yang lebih baik dari penelitian sebelumnya. Penelitian ini menggunakan dataset yang disediakan oleh Kaggle dengan jumlah 100 data *training* dan 25 data *testing*. Oleh sebab itu, peneliti melakukan augmentasi data agar dapat memperbanyak dataset. Untuk model modifikasi VGG 16 yang dibuat untuk melakukan klasifikasi, mendapatkan hasil evaluasi rata-rata akurasi sebesar 81%, nilai *recall* sebesar 76%, dan nilai *precision* sebesar 81% untuk fase *training* dan untuk fase validasi, akurasi sebesar 85%, nilai *recall* sebesar 80%, dan nilai *precision* sebesar 84%. Jadi dengan model modifikasi VGG 16 dapat disimpulkan bahwa model mampu memprediksi rempah-rempah lebih baik dari model Alexnet selain itu jika dibandingkan dengan penelitian sebelumnya, model modifikasi VGG pada penelitian ini memiliki nilai evaluasi akurasi dan *loss* yang lebih baik. Untuk penelitian selanjutnya, perlu dikembangkan datasets rempah yang lebih baik dan beragam dari sisi rempah-rempah yang digunakan, jumlah citra pada masing-masing kelasnya, dan citra yang memiliki keadaan-keadaan tertentu seperti citra pada saat rempah masih baru di ambil dan lainnya. Untuk model yang digunakan, masih perlu pengembangan arsitektur yang tepat agar model mampu mengklasifikasi rempah-rempah yang memiliki kelas lebih banyak dibandingkan dengan dataset yang digunakan pada penelitian ini baik itu membuat arsitektur dengan menambahkan convolution layer ataupun menggunakan metode image processing lainnya untuk memodifikasi citra rempah agak lebih mudah di klasifikasi.

#### REFERENSI

- [1] A. W. Helmalia, P. Putrid, and A. Dirpan, "Potensi Rempah-Rempah Tradisional Sebagai Sumber Antioksidan Alami Untuk Bahan Baku Pangan Fungsional," *Canrea J. Food Technol. Nutr. Culin. J.*, vol. 2, no. 1, pp. 26–31, 2019.
- [2] Y. Robi, S. M. Kartikawati, and . Muflihati, "Etnobotani Rempah Tradisional di Desa Empoto Kabupaten Sanggau Kalimantan Barat," *J. Hutan Lestari*, vol. 7, no. 1, pp. 130–142, 2019.
- [3] K.V. Peter, *Handbook of Herbs and Spices*. Elsevier Science, 2012.
- [4] Liputan6dotcom, "Sulit Bedakan Rempah-Rempah seperti Jahe, Lengkuas, Kunyit, dan Kencur? Begini Caranya," <https://www.liputan6.com/>, 2018. .
- [5] I. Wulandari, H. Yasin, and T. Widiari, "Klasifikasi Citra Digital Bumbu dan Rempah Dengan Algoritma Convolutional Neural Network (Cnn)," *J. Gaussian*, vol. 9, no. 3, pp. 273–282, 2020.
- [6] D. C. Khrisne and I. A. Suyadnya, "Indonesian Herbs and Spices Recognition using Smaller VGGNet-like Network," *2018 Int. Conf. Smart Green Technol. Electr. Inf. Syst.*, vol. 4, pp. 221–224, 2018.
- [7] Kaharuddin, Kusrini, and E. T. Luthfi, "Klasifikasi Jenis Rempah-Rempah Berdasarkan Fitur Warna Rgb dan Tekstur Menggunakan Algoritma K-Nearest Neighbor," *J. Inf. Interaktif*, vol. 4, no. 1, pp. 17–22, 2019.
- [8] R. Xin, J. Zhang, and Y. Shao, "Complex network classification with convolutional neural network," *Tsinghua Sci. Technol.*, vol. 25, no. 4, pp. 447–457, Aug. 2020.
- [9] V. Atliha and D. Sesok, "Comparison of VGG and ResNet used as Encoders for Image Captioning," *2020 IEEE Open Conf. Electr. Electron. Inf. Sci. eStream 2020 - Proc.*, pp. 1–4, 2020.
- [10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, *Handbook of Approximation Algorithms and Metaheuristics*. Chapman and Hall/CRC, 2007.

- [11] “Rempah | Kaggle.” <https://www.kaggle.com/awaltry/rempah> (accessed Oct. 05, 2021).
- [12] R. Rismiyati and A. Luthfiarta, “VGG16 Transfer Learning Architecture for Salak Fruit Quality Classification,” *Telematika*, vol. 18, no. 1, p. 37, 2021.
- [13] Y. Zhou, H. Ni, F. Ren, and X. Kang, “Face and gender recognition system based on convolutional neural networks,” *Proc. 2019 IEEE Int. Conf. Mechatronics Autom. ICMA 2019*, pp. 1091–1095, 2019.
- [14] N. Tang, H. Liu, K. Yue, W. Li, and X. Yue, “Automatic classification for corneal ulcer using a modified VGG network,” *Proc. - 2020 Int. Conf. Artif. Intell. Comput. Eng. ICAICE 2020*, pp. 120–123, 2020.
- [15] Y. Q. Xuan Liu, Mingmin Chi, Yunfeng Zhang, “Classifying High Resolution Remote Sensing Images By Fine-Tuned Vgg Deep Networks,” pp. 7141–7144, 2018.
- [16] A. A. Almisreb, N. Jamil, and N. M. Din, “Utilizing AlexNet Deep Transfer Learning for Ear Recognition,” *Proc. - 2018 4th Int. Conf. Inf. Retr. Knowl. Manag. Diving into Data Sci. CAMP 2018*, pp. 8–12, 2018.
- [17] C. Dong, Z. Zhang, J. Yue, and L. Zhou, “Classification of strawberry diseases and pests by improved AlexNet deep learning networks,” *2021 13th Int. Conf. Adv. Comput. Intell. ICACI 2021*, pp. 359–364, 2021.
- [18] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, Dec. 2014, pp. 1–15.